

Music sound quality assessment in bimodal cochlear implant users – toward improved hearing aid fitting

Khaled H. A. Abdellatif¹, Verena Müller², Martin Walger^{1,2}, Hartmut Meister¹

¹ Jean-Uhrmacher-Institute for Clinical ENT-Research, University of Cologne, Geibelstr. 29-31, D-50931 Cologne, Germany.

² Department of Otorhinolaryngology, Head and Neck Surgery, Medical Faculty, University of Cologne, Kerpenerstraße 62, D-50937 Cologne, Germany.

Abstract: The use of cochlear implants (CI) is the common treatment to counteract severe-profound hearing loss. CI often allow reasonable speech understanding but are generally limited in terms of spectro-temporal resolution. Many CI recipients use a hearing aid (HA) on the non-implanted ear (so called bimodal patients) that complements the electrical stimulation of the CI by providing acoustical fine structure information. Amongst others, this might be especially beneficial for music appraisal. However, it is not clear how the HA should be fitted in conjunction with the CI in order to provide the best sound quality.

The purpose of this study was to vary different fitting parameters of the HA systematically and to determine the resulting music sound quality. To this end, a master hearing aid was used in order to have controlled access to the different parameters. Three different music excerpts (pop with and without vocals, classic) were presented to 13 bimodally fitted CI users and the sound quality was assessed using the “multiple-stimulus with hidden reference and anchor” (MUSHRA) test. Since the goal was to gain information on possible improvement of fitting, a complete retest was performed and individual reliability of the assessments was determined by the eGauge method.

It could be shown that most of the listeners were able to provide reliable sound quality judgements. In terms of HA fitting it turned out that changing compression and varying low frequency gain had a significant effect compared to a standard prescriptive fitting based on DSL v5.0. However, especially the effect of changing low frequency gain depended on the music excerpts presented, since pop music with vocals revealed the largest effect.

The study gives evidence that hearing aid fitting can be optimized relative to a standard prescriptive rule in order to improve music sound quality in bimodal CI users. Moreover, the method might be helpful in a clinical setting to determine the best fitting based on individual results.

Key words: Sound quality, cochlear implant, hearing aid, bimodal fitting, MUSHRA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Corresponding author: hartmut.meister@uni-koeln.de

Introduction

Cochlear implantation allows partially restoring hearing in severe-to-profoundly hearing impaired individuals. The primary goal of cochlear implant (CI) provision is to improve speech recognition and hence to enable the listeners to participate in verbal communication. Indeed, many CI users reveal reasonable speech understanding at least in quiet communication situations (Clark, 2015). However, CI-mediated listening is far from the abilities a healthy auditory system allows for. Besides differences in dynamic range (see Moore, 2022) this is especially due to the fact that CI for example typically provide a restricted spectro-temporal resolution, because of the limited number of electrodes that can be used to stimulate the auditory nerve (e.g., Friesen et al., 2001). Additionally, interaction between the channels blurs the frequency mapping and temporal pitch cues are only of limited use. As a consequence, a CI mainly transmits envelope information as one of the features important for speech recognition (e.g., Rubinstein, 2004). However, the lack of spectro-temporal fine-structure may have consequences for other listening tasks.

In recent years, an increasing research effort has been directed towards CI users' perception of non-speech sounds, especially music. Music is based on a number of different dimensions, such as rhythm, melody, harmony, and timbre. In terms of CI-mediated listening, several limitations arise in these dimensions (overview in McDermott, 2004). Due to the pulsatile stimulation patterns of CI, temporal features of music such as rhythm are generally relatively well preserved. In contrast, the limited number of spectral channels restricts melodic tune identification (Looi et al. 2004), melody recognition (Gfeller et al. 1997) and perception of timbre (Gfeller et al. 2002). The latter is related to different instruments and thus helps distinguishing multiple signals in music. It is also significantly connected to sound quality, which is itself a determinant of consumer satisfaction (e.g., Wong et al., 2003). Though there is no generally accepted definition of sound quality, it can be seen as a construct that is constituted by the acoustic properties of a stimulus and by the perceptual assessment, for instance in terms of accuracy or fidelity. Moreover, expectations might play a role with the assessment of sound quality. As such, sound quality is a highly subjective domain.

Given the apparent limitations of electric stimulation on music appraisal, a number of measures for improving sound quality have been discussed. A way to tone down the negative consequences of limited spectro-temporal fine structure is the use of a hearing aid (HA) on the ear contralateral to the CI (bimodal stimulation) or even in combination with the CI on the implanted ear (hybrid stimulation). In these cases the HA mainly transmits sounds from the low-frequency range where residual hearing is typically largest. Many CI recipients use bimodal stimulation that has positive consequences on speech perception (e.g., Ilg et al., 2014). Beyond that, also music perception and sound quality may be improved (Sucher & McDermott, 2009).

However, it is still unclear how the HA should be fitted in order to achieve the best sound quality in bimodal patients. HAs are typically adjusted using generic fitting rules such as DSL v5.0 (desired sensation level, Scollie et al., 2005), NAL-NL2 (Keidser et al., 2011), or manufacturer-specific approaches (e.g., Vroegop et al., 2019). These rules use patient-specific data typically from the pure-tone audiogram as input and calculate prescriptive gain as a function of frequency, level, and hearing loss for long-term average speech spectrum signals. However, there is currently no generally accepted fitting rule for a HA in combination with the CI in order to maximize perception. In the framework of a survey on audiologists, Scherf & Arnold (2014) found many of them reporting that the HA is not re-fitted when used in conjunction with a CI. This often has practical reasons since CI and HA provision may take place at different institutions. Others reported that at least a loudness balancing between the two devices is performed, as suggested by Ching et al. (2004). However, a

structured procedure to consider the apparent differences between the CI and HA, which have to be taken into account for bimodal hearing, obviously does not exist yet.

Recently, Vroegop et al. (2019) tapped into this issue by comparing three different approaches of hearing aid fitting in the framework of bimodal listening. They considered two different fitting formulas as well as a loudness balancing based on broadband and narrowband signals. The different procedures were assessed by means of localization and speech recognition but in general, no clear differences were found. This could be interpreted as once sufficient gain is provided by the HA, further fine-tuning in conjunction with the CI does not significantly increase the outcome. However, it must be kept in mind that the study focused on performance measures and that the findings might be different when subjective criteria, such as (music) sound quality, are considered.

The present study aims at addressing this aspect by systematically assessing the effects of different HA-adjustments on sound quality ratings in bimodal listening. Here, the rationale is to manipulate only the HA and use the CI in its regular, optimally fitted condition. In order to keep the factors controllable, a “master hearing aid” (openMHA, Grimm et al., 2006) was used instead of real hearing aids. Three different excerpts of music were considered and transmission via the HA was manipulated in terms of frequency range, compression as well as low frequency gain. The ultimate goal of this study is to give information for a possible modification of HA fitting procedures in terms of improving sound quality. As the outcome is based on subjective ratings, reliability of the individuals’ quality estimations is a crucial point to reach this goal and was thus an important element of the study.

Methods

Subjects

Thirteen bimodally fitted listeners (3 female and 10 male) with a mean age of 55 years (range 28–71 years, SD=17.2) were recruited in this study (see Table 1 for details). They were all provided with Cochlear™ sound processors CP910 (n=1) and CP1000 (n=12). This restricted choice of devices was due to keep the methodological and especially the technical variability small. Residual hearing of the non-implanted ear was mainly present in the frequency range up to 1 kHz. The subjects were chosen to have a hearing loss of not larger than 85 dB HL in this frequency range as no benefit in music perception is expected for thresholds exceeding this value (El Fata et al., 2009). The corresponding low frequency average (250-1000 Hz) was between 38.3 and 83.7 dB HL.

All participants were native German speakers and had normal or corrected-to-normal vision. Prior to the experiment, they were given detailed information about the study and informed consent was obtained. Participants were reimbursed with € 10,-/h. The study protocol was approved by the local ethics committee (reference 18-383).

Setup

The setup is based on a computer (Samsung Ultrabook Model: NP530U4E running Windows 8) where the stimuli are generated and the participant’s responses are collected. Since the rationale was to modify hearing aid fitting while leaving the cochlear implant untouched, the original stimulus was transmitted to the CI. This was accomplished by connecting a soundcard (MAYA44USB, Ego Systems Inc. and ESI Audiotechnik GmbH) to the Cochlear™ Wireless TV Streamer and transmitting via bluetooth to the sound processor. This transmission line caused a delay of 40 ms relative to the

acoustic stimulation of the non-implanted ear, which was compensated for by delaying the contralateral stimulus accordingly.

The acoustic stimulation was conducted via the open Master Hearing Aid (openMHA, Grimm et al., 2006). Using the openMHA instead of a real hearing aid was essential for standardization of signal transmission and considers basic aspects such as frequency-dependent amplification and dynamic-range compression (DRC). A 10-channel multiband dynamic compression with center-frequencies ranging from 250 Hz to 8000 Hz was implemented by the openMHA. Based on the individual pure-tone audiogram, frequency-gain shaping and dynamic range compression were prescribed for each listener using DSL v5.0 gains for 50, 65, and 80 dB SPL input levels using the Otometrics OTOsuite software (Natus Medical Denmark ApS). Linear inter-/extrapolation was used to obtain gains for input levels ranging from 0 to 120 dB SPL in steps of 1 dB. Gains were limited so that the resulting predicted real-ear levels did not exceed the OSPL90 values provided by DSL v5.0. Gain values were manually inputted as a gain table into the openMHA software. The fast-acting compression attack and release times were 0.02 and 0.1 seconds, respectively.

The audio output of the soundcard was connected to a headphone amplifier (Superlux HA3D, Superlux Enterprise Development (Shanghai) CO., LTD) and transmitted on a hearing aid receiver (SureFit2 (HP), ReSoundGN) coupled to the ear canal by a tympanometry tip and enclosed in an impression material. This closed fitting ensured that no acoustic leakage occurs, which may severely reduce low-frequency energy.

ID	Gender	Age range [yrs]	CI experience [yrs]	HA experience [yrs]	LF PTA [dB HL]	LF gain stepsize [dB]	Reliab. classic	Reliab. Pop with vocals	Reliab. Pop no vocals
S01	f	71-75	5	20	70.0	5.1	yes	yes	yes
S02	m	56-60	4	4	63.0	7.1	no	yes	yes
S03	m	51-55	1.5	3	43.7	6.0	yes	yes	yes
S04	m	66-70	1.6	55	82.3	7.2	yes	yes	no
S05	m	51-55	1.1	27	39.3	4.5	yes	yes	yes
S06	m	26-30	1.1	20	83.7	5.8	yes	yes	yes
S07	f	76-80	1.0	11	38.3	12.6	no	no	no
S08	m	26-30	6.0	16	76.0	7.0	yes	yes	yes
S09	m	66-70	1.0	6	81.0	11.3	yes	no	yes
S10	m	66-70	0.3	16	44.7	6.3	yes	yes	yes
S11	f	36-40	0.6	16	55.3	8.0	yes	yes	yes
S12	m	36-40	0.3	31	80.0	6.2	yes	no	yes
S13	m	61-65	2	26	53.0	10.7	yes	no	yes

Table 1: Subject characteristics. Low frequency pure-tone average (LF PTA) was calculated across 250, 500, and 1000 Hz. LF gain stepsize (averaged across the three music excerpts) displays the value derived from the just noticeable differences (jnds) and taken as stepsize 1. The three right columns show the reliability (Reliab.) of the data as assessed by the eGauge method. For reasons of anonymity the age range is given.

Music stimuli

Three excerpts of different music pieces were included in the assessment (duration 14 seconds each). Pop music with vocals was represented by an excerpt of Namika's "Je ne parle pas francais" including a standard instrumentation (guitar, keyboard, bass, and drums), and a female vocal. The same

excerpt was considered as a pure instrumental version in order to examine the effect of vocals on sound quality. Moreover, classic music was covered by an excerpt of Johann Strauss' "The blue danube" played by a full orchestra representing a complex instrumentation. These excerpts were chosen to give different genres and different acoustic characteristics (i.e., rhythm, timbre, complexity, vocals). They were also chosen in order to represent a reasonable amount of acoustic information in the low frequency range (i.e., below 1 kHz) where residual hearing typically plays the most important role. To this end, the spectral centroid was calculated for each music piece using the matlab MIRtoolbox (Lartillot and Toivainen 2007). The average spectral centroid across the duration of the music pieces was 683 Hz for classic, 613.8 Hz for pop with vocals and 383 Hz for pop without vocals. Shorter parts of these three recordings (duration 2.5 seconds) were additionally used to assess just noticeable differences (jnds) regarding changes in low-frequency gain. These additional measurements were performed to ensure that the modifications of the stimuli in the MUSHRA assessment (see below) were actually perceivable by the listeners.

Sound quality assessment

The "multiple-stimulus with hidden reference and anchor" (MUSHRA) test was used for sound quality assessment (see recommendation ITU-R BS.1534-3). This method allows simultaneous comparison and rating of different stimuli. Typically, the stimuli are compared against a known reference that is chosen to have the best quality. However, in our case the reference (i.e., fitting based on DSL v5.0) was only presented in a hidden form, since it must not necessarily present optimal sound quality. The anchor, which should represent poor quality, was generated by a noise vocoder (Gaudrain and Başkent, 2015). The MATLAB code is available online (see Gaudrain, 2016). The signal was filtered into two bands having equally spaced boundaries based on a 35-mm basilar membrane distance (Greenwood, 1990) across a frequency range between 0.2 and 20 kHz. Then the output of each band was half-wave rectified and low-pass filtered at 250 Hz to extract the temporal envelope. This was then multiplied by a wide-band noise carrier, and the resulting signal was summed across the bands. Rating was performed on a 100 point scale with semantic labelling of sound quality (i.e., "bad" (range 0-20 pts), "poor" (20-40 pts), "fair" (40-60 pts), "good" (60-80 pts), and "excellent" (80-100 pts). This way, the "mean opinion scores" (MOS) were derived. The MUSHRA "drag & drop" version (Völker et al., 2018), which represents a more intuitive user-interface, was applied for the sound quality assessment. This version allows activating the stimulus by clicking on the corresponding button (always labelled A-G for the different stimuli). The next step is to drag the button into the rating field and to drop it onto the appropriate rating position. All stimuli could be played and dragged as often as desired until a final decision about the sound quality was reached.

Procedures

After the participants were instructed about the study rationale informed consent was obtained. The sound quality assessment based on MUSHRA was conducted for two conditions, a "general" and a "fitting" condition, each with the DSL-based gain adjustments as the hidden reference presented twice in each condition. The "general condition" considered more basic aspects of bimodal listening, such as sound quality for each device alone, linear vs. non-linear amplification as well as overlap vs. non-overlap of the frequencies of the cochlear implant and the hearing aid (i.e., $HA \leq 1$ kHz, $CI > 1$ kHz). In the latter case, not the original stimulus was sent to the CI but rather a low-pass filtered version with a cutoff-frequency of 1 kHz, approximating the upper limit of the main frequency range transmitted by the openMHA. The no-overlap condition was motivated by the fact, that some

literature pointed to a decrease in music perception when CI and HA share the same frequency region. (e.g., Crew et al., 2016).

The “fitting condition” considered gain manipulations of the HA in the low frequency range ≤ 1 kHz. Gain was either decreased or increased by one or two individual steps. Individual stepsize was used to ensure that gain manipulations were always perceivable. Therefore just noticeable differences (jnds) were assessed in a 3I3AFC paradigm (Ewert, 2013) presented via the openMHA. Jnds associated with 75% correct discrimination were determined for changes in low-frequency gain. Stepsize 1 was set to 1.5 jnd, consequently approximating differences that were small but detectable. Individual stepsize averaged across the three music excerpts is shown in Table 1. Stepsize 2 represented the double value of step 1. For large jnds the stepsize might cause clipping. However, in this case stepsize 2 was limited to the maximum gain value before any distortions occurred.

Prior to the sound quality ratings individual intensity adjustments to reach the most comfortable loudness level were carried out for each music piece. This was done for the CI and the HA alone as well as for the combination of both devices. After loudness adjustment, half of the listeners began quality assessment in the general condition and half in the fitting condition. The presentation of the different music excerpts was pseudo-randomized. After a break of 15 minutes a complete retest was performed in order to check reliability. In total each listener gave 84 assessments (3 music pieces x 7 processing x 2 conditions x 2 tests) using the MUSHRA procedure. Total testing time was about 2 hours.

Reliability of the ratings was assessed by eGauge (Lorho et al., 2010). This method calculates the reliability using a non-parametric permutation test (Dijksterhuis & Heiser, 1995) as a test of significance. The permutation test is computed using several iterations per subject and defines the noise floor of the performance for reliability metrics. Below that level, the performance is equivalent to random ratings that could degrade quality of data. Consequently, sound quality assessments where the listeners did not fulfil this criterion were not considered for further analysis (see Table 1).

Statistical analysis

Data were subjected to a one-way repeated-measures analysis of variance (rmANOVA) with processing condition as within-subject factor. Greenhouse-Geisser correction was applied if the assumption of sphericity was violated. Paired comparisons with Bonferroni correction were conducted as post hoc analyses. The p-value was set to 0.05. Analyses were performed with IBM SPSS statistics v. 27.

Results

General condition

Figure 1 shows the results of the “general condition” for the excerpt of classic music (panel A), pop with vocals (panel B), and pop without vocals (panel C). Panels A and B show relatively similar patterns with the DSL-reference revealing the best and the anchor revealing the worst sound quality. However, panel C appears to be different since all conditions show relatively similar outcome and the variability of the MOS is higher than for the other two music pieces.

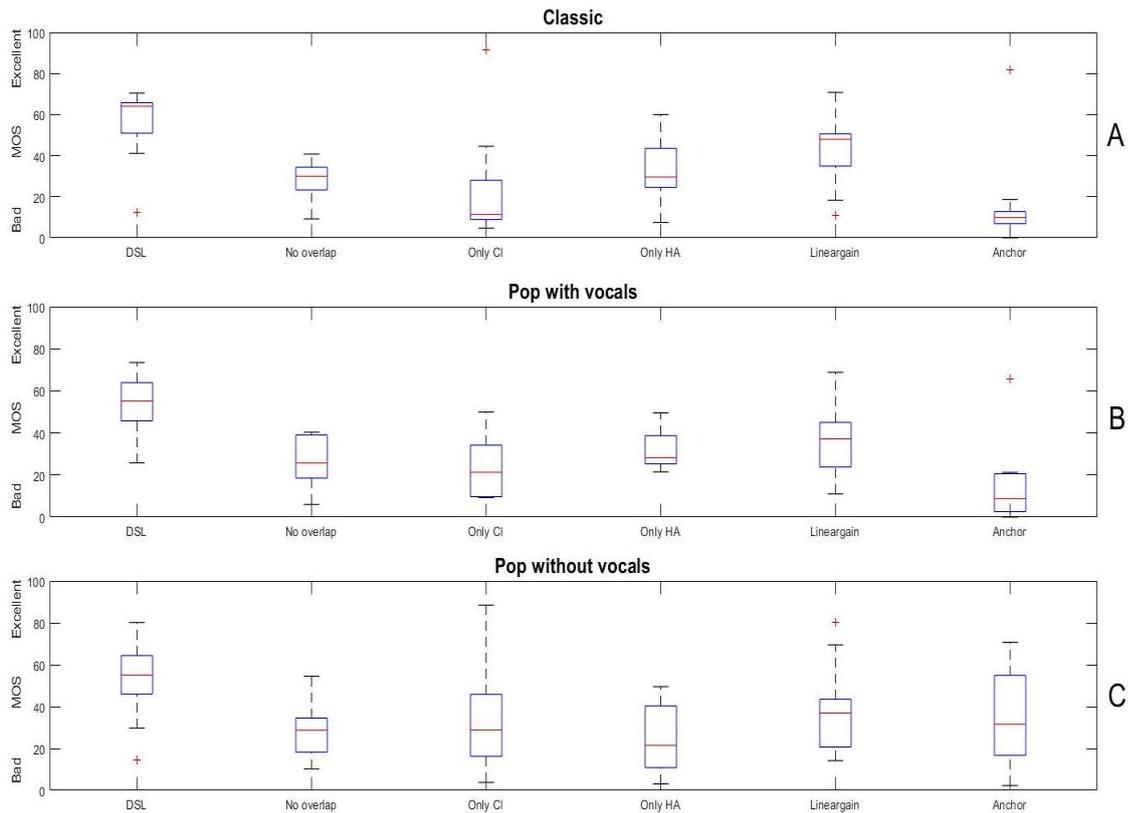


Figure 1: Mean opinion scores (MOS) for the excerpt of classic music (A), pop music with vocals (B) and pop without vocals (C). The conditions displayed on the x-axis are given in the text.

A rmANOVA on the MOS-data of the *classic music excerpt* with processing condition as the independent variable shows a significant main effect of processing ($F(1.721, 5)=9.6$, $p=0.002$, $\eta_p^2=0.49$). Post hoc pairwise comparisons with Bonferroni-correction reveal a significantly higher MOS for DSL compared to all other processing conditions ($p\leq 0.031$). Moreover, MOS for the anchor stimulus reveals significantly lower sound quality than both, DSL ($p=0.011$) and linear fitting ($p=0.003$).

A rmANOVA on the MOS-data of the *pop music excerpt* with processing condition as the independent variable also shows a significant main effect of processing ($F(2.562, 5)=8.375$, $p=0.001$, $\eta_p^2=0.511$). Post hoc pairwise comparisons with Bonferroni-correction reveal a significantly higher MOS for DSL compared to the CI alone ($p=0.010$) and the HA alone ($p=0.025$) whereas DSL is not significantly different from the no-overlap setting ($p=0.063$) and the linear gain condition ($p=0.081$), based on the chosen significance level. The anchor stimulus is rated significantly poorer than the DSL-reference ($p=0.005$).

Lastly, a rmANOVA on the MOS-data of the *pop music excerpt without vocals* with processing condition as the independent variable again shows a significant main effect of processing ($F(1, 5)=4.168$, $p=0.003$, $\eta_p^2=0.294$). Post hoc pairwise comparisons with Bonferroni-correction reveal a significantly higher MOS for DSL compared to HA alone ($p=0.010$) and the no-overlap condition ($p=0.015$). However, no further significant differences can be found. Notably, the anchor stimulus is not rated significantly worse than any of the other processing conditions.

Taken together, both the classic and pop music excerpt show the highest sound quality for the DSL-reference and the lowest sound quality for the anchor stimulus, as intended. Listening with the CI

and the HA alone yields significantly worse sound quality than the combination of both devices. However, this does not apply for the no overlap condition that yields MOS in the range of the single devices. The linear processing condition is also not significantly different than the two devices separately but reveals higher MOS at least in absolute terms.

The pop music excerpt without vocals appears to cause somewhat different results, since the conditions revealed more similar results and the anchor did not yield the intended low sound quality. In fact, despite an absolute difference of about 15 points the anchor was not rated significantly worse than the DSL reference.

Fitting condition

Figure 2 shows the results of the “fitting condition” for the excerpt of classic music (panel A), pop with vocals (panel B), and pop without vocals (panel C). Again, the DSL-reference mostly yields highest sound quality but changing low-frequency gain appears to affect the ratings.

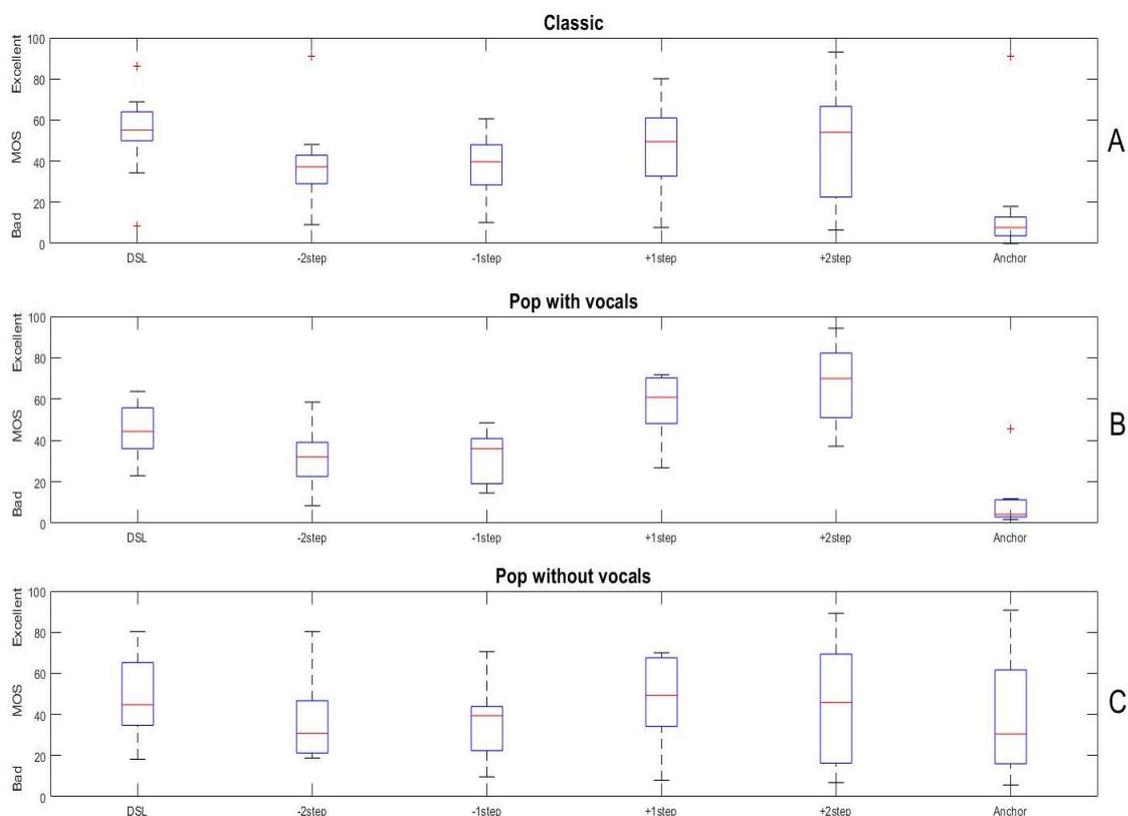


Figure 2: Mean opinion scores (MOS) for the excerpt of classic music (A), pop music with vocals (B) and pop without vocals (C). The conditions displayed on the x-axis are given in the text.

A rmANOVA on the MOS-data of the *classic music excerpt* with processing condition as the independent variable shows a significant main effect of processing ($F(2.306, 5)=6.2$, $p=0.005$, $\eta_p^2=0.383$). Post hoc pairwise comparisons with Bonferroni-correction reveal a significantly higher MOS for DSL compared to the anchor stimulus ($p=0.004$) as well as the gain reduction of one ($p=0.003$) and of two steps ($p=0.007$).

A rmANOVA on the MOS-data of the *pop music excerpt* with processing condition as the independent variable shows a significant main effect of processing ($F(1, 5)=27.715$, $p<0.001$, $\eta_p^2=0.776$). Post hoc pairwise comparisons with Bonferroni-correction reveal a significantly higher MOS for DSL compared to the anchor stimulus ($p=0.004$) and the gain reduction of two steps ($p=0.018$). However, a significantly higher MOS for a gain increase of 2 steps compared to the DSL-reference can also be observed ($p=0.017$).

In contrast, a rmANOVA on the MOS-data of the *pop music excerpt without vocals* with processing condition as the independent variable does not show a significant effect of processing ($F(1.641, 5)=0.649$, $p=0.506$, $\eta_p^2=0.061$).

Basically, this analysis revealed that manipulating low frequency gain affects sound quality assessments. Still, the effect seems to depend on the piece of music presented. Whereas for the classic excerpt only a decrease in sound quality with decreasing gain relative to the DSL-reference could be found, a better sound quality with increasing gain was additionally observed for the pop music excerpt with vocals. However, when this piece of music was presented in an instrumental version, no significant effect of gain manipulation could be shown.

Discussion

This study focused on music sound quality in bimodal CI users by manipulating the settings of a (master) hearing aid and determining the corresponding effects by using a multiple-stimulus (MUSHRA) method. The overall goal of the study was to suggest possible modifications in hearing aid fitting that might improve subjectively perceived sound quality.

The study considered two different conditions. One condition tapped into more general aspects of bimodal fitting such as the sound quality perceived with either device alone, the overlap of the frequency regions transmitted by the CI and the HA as well as a linear amplification, compared to the compressive standard DSL fitting. The second condition addressed a more direct fitting aspect, namely the decrease or increase of low frequency gain of the hearing aid. The fundamental study concept was that any manipulation was performed on the hearing aid while keeping the settings of the cochlear implant as used by the individual listener.

Three different music excerpts were selected. Their choice was based on different considerations. First, in order to reflect manipulations of the HA that typically provides low frequency gain in bimodal CI users, a significant proportion of the spectrum should lie in this frequency area. This was ensured by determining the spectral centroid of the stimuli that was below 1 kHz in all cases. Second, the music excerpts should represent differently complex stimuli by taking a piece with a rich instrumentation (classic), and a more simple instrumentation with and without vocals into account (pop). The fact that vocals appear to play a unique role for CI recipients (e.g., Riley et al., 2018) was taken into account by using the same excerpt of a single song, once with and once without vocals.

Sound quality judgements were assessed using the multiple-stimulus with hidden reference and anchor test (MUSHRA), typically denoted as CI-MUSHRA when used with cochlear implant recipients (Roy et al., 2015). We used the “drag & drop” version of the MUSHRA that has practical advantages as it presents a more intuitive user-interface (Völker et al., 2018).

General condition

The general condition revealed that the reference fitting based on DSL provided the highest sound quality. Using only the CI or the hearing aid significantly decreased MOS, as expected. There was no significant difference between using the two devices in isolation. Since MOS coarsely doubled in the combined reference condition this might give evidence that the CI and the HA largely present complementary sound cues (Ching et al., 2006). Hence, both the envelope information over a broader frequency range, predominantly transmitted by the CI, and the acoustic fine structure cues in the low-frequency area transmitted by the HA contribute to music sound quality.

The no-overlap condition was considered to control whether sharing the same frequency range (i.e., below 1 kHz) is detrimental, as discussed in the literature. Though not conclusive, some studies reveal negative effects of overlap and non-optimal frequency-to-place coding with bimodal stimulation (e.g., Francart & McDermott, 2013, Crew et al., 2016). However, we found that in the “no-overlap” condition sound quality clearly decreased. This could be due to several reasons. First, it is unclear whether a changed frequency allocation of the CI caused by the low-pass filtered stimuli impacted ratings negatively. Another factor could have been that the “no-overlap” condition disturbs the impression of “stereo listening”, which is important for sound quality, as recently shown by Landsberger et al. (2020).

Finally, we also included a condition based on DSL, but with a linear sound processing instead of the dynamic compression. This condition somewhat decreased perceived sound quality, although the difference to the DSL-setting was not statistically significant for all music excerpts. The influence of compression on sound quality ratings is inconclusive. Some studies have shown that compressive is preferred over linear fitting. For instance, Gilbert et al. (2021) manipulated the “maplaw” (i.e., the mapping of the acoustic input to the electric dynamic range) as well as the automatic gain control (AGC) in Med-El CI-processors. Whereas AGC had only subtle effects on perceived sound quality, higher maplaw (i.e., more compression) was associated with better quality ratings. Subjective feedback of the listeners revealed better audibility for soft sounds and lyrics and basically a “fuller” impression of the sound as a possible reasons. In contrast, Kirchberger & Russo (2016) showed in hearing aid users with predominantly moderate hearing loss that linear was preferred over compressive processing with different compression ratios. However, results from hearing aid users might not be directly comparable with bimodal cochlear implant users as different frequency regions are transmitted and the tradeoff between increased audibility of soft sounds and possible distortions caused by compression may be different.

These results apply to the classic and pop music excerpt with vocals in a similar manner. However, notably the pop music piece without vocals showed somewhat unusual results in that differences between the various conditions were clearly smaller. Particularly the anchor stimulus revealed a relatively high rating that was not seen for the two other music excerpts. This suggests that for the sparsely instrumented stimulus without vocals different cues are at work. We speculate that mainly rhythmic aspects played a role here, which were less sensitive to the signal manipulations applied – including the noise vocoding performed to generate the anchor stimulus.

Fitting condition

The fitting condition specifically addressed changes of low-frequency gain. Low frequencies play an important role for sound quality, however, this frequency area might be critical for CI users due to a potential lack of stimulation in the apical cochlear region (Roy et al., 2012, Landwehr et al., 2013; Roy

et al., 2015). Hence, complementary transmission of this information by the hearing aid might be especially crucial.

The fitting condition showed significant effects of low frequency gain manipulations. Compared to the DSL reference, decreasing the low frequency gain yielded a significantly lower sound quality in both, the classic and pop music excerpt. We assume that reducing low frequencies makes the sound “thinner”, hence negatively affecting perceived quality. Moreover, increasing low frequency gain significantly increased sound quality relative to DSL, but only for pop music with vocals. This is consistent with the proposed special role of vocals for sound quality in listeners with CI (Riley et al., 2018). For instance, Gfeller et al. (2008) stated that “when musical stimuli have lyrics, some factors that have been shown to influence speech perception also influence accuracy of music perception”. Moreover, Buyens et al. (2014) asked normal-hearing listeners and cochlear implant recipients to mix the tracks of different music pieces in order to generate a version that sounded most enjoyable. The CI users generated mixes that showed a relatively more intensive adjustment of the voices. In line with this is the recent study by Limb et al. (2022) who showed that music was preferred by CI users when the voice of the sound was amplified by 9 dB – a value that is in the range of the gain-steps adjusted in the present study (see Table 1).

The improvement of sound quality due to increased low-frequency gain may not necessarily be restricted to cochlear implant recipients. Vaisberg et al. (2021) asked listeners with mild-moderate hearing loss to adjust gain settings as individually preferred. Similar to our study they used the openMHA as a hearing aid simulation. In comparison to the standard DSL-fitting, largest adjustment of low-frequency gain in pop music followed by classical music were found. The authors concluded that a standard fitting might be appropriate in terms of speech recognition but not necessarily when listening to music.

Taken together our study provides some evidence for an improvement in music sound quality in bimodally fitted CI users by manipulating the low-frequency acoustic transmission via the hearing aid. First, linear amplification appeared to be inferior to dynamic compression as applied based on the standard DSL- fitting. This was found for all of the three music excerpts though the difference was not always statistically significant – possibly due to the large inter-individual variance in sound quality ratings. Secondly, sufficient low frequency gain – as proposed by DSL – is important. Otherwise, sound quality is compromised, as was shown with decreasing gain for both pop with vocals and the classic piece. Furthermore, sound quality was improved by increasing the low frequency gain in the pop music excerpt with vocals. Consistent with other research, this supports the extra role of vocals for music perception with CI. In contrast, gain manipulations in sparsely instrumented music, such as the piece of pop without vocals, show a much smaller effect. We suspect that this is due to the superior role of rhythmic cues. Thus, when trying to improve sound quality by modifying hearing aid fitting in bimodal cochlear implant recipients, different types of music need to be considered.

Strengths and limitations of the study

This study used the openMHA in order to provide the acoustic path to the listener. We opted for the openMHA instead of the individual hearing aid of the listener because it offers straightforward access to the relevant parameters (in our case compression and frequency dependent gain) and manipulations can be performed in an arbitrary manner. This would not have been readily feasible in real hearing aids where (unknown) signal processing algorithms may have been at work. For a similar reason we decided to take only two types of speech processors from the same manufacturer into account, which also helped to keep methodological variability low. Another strength of the study is

the assessment of the reliability of the individuals' answers. Using eGauge it could be shown that most but not all of the subjects provide reliable ratings. To the best of our knowledge, this is the first study that explicitly determined the reliability of sound quality ratings in CI users.

A limitation is the small study sample that was mainly caused by the restricted choice of processor types, as explained above. This made it particularly difficult to perform a subgroup-analysis, e.g. based on the residual hearing of the participants. Another restriction was the choice of a relatively small sample of music excerpts, which however were specifically selected, and the use of the openMHA instead of a real hearing aid, which might limit transfer of the results to real life conditions. This latter point is currently addressed in a follow-up study. Last not least, as an acute trial this study does not consider long-term adaptation effects.

Acknowledgements

Supported by Cochlear Ltd. (Reference IIR-1934). We are grateful to the participants for their time and dedication. Thanks to Horst Hessel for fruitful discussions and Christoph Völker for providing the MUSHRA test.

Results of this study were presented at the 15th Congress of European Federation of Audiology Societies 2021, 20. May 2021, and the 3rd Music and Cochlear Implant Symposium, Cambridge, 15./16. September 2021. The study has been registered in DRKS (Deutsches Studienregister), DRKS00016778.

References

- Buyens, W., van Dijk, B., Moonen, M., & Wouters, J. (2014). Music mixing preferences of cochlear implant recipients: A pilot study. *International journal of audiology*, 53(5), 294-301.
- Ching T.Y., Hill M., Dillon H., van Wanrooy E. Fitting and evaluating a hearing aid for recipients of a unilateral cochlear implant: the NAL approach, Part 1. Hearing aid prescription, adjustment, and evaluation. *Hearing Rev* 2004;14–58.
- Ching, T. Y., Incerti, P., Hill, M., & Van Wanrooy, E. (2006). An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices. *Audiology and Neurotology*, 11(Suppl. 1), 6-11.
- Clark, G.M. (2015). The multi-channel cochlear implant: multi-disciplinary development of electrical stimulation of the cochlea and the resulting clinical benefit. *Hear Res.* 322, 4–13. doi: 10.1016/j.heares.2014.08.002
- Crew, J. D., Galvin III, J. J., & Fu, Q. J. (2016). Perception of sung speech in bimodal cochlear implant users. *Trends in hearing*, 20, 2331216516669329.
- Dijksterhuis, G.B. and Heiser, W.J. (1995). The role of permutation tests in exploratory multivariate data analysis, *Food quality and preference* 6, 263-270.
- El Fata, F., James, C. J., Laborde, M. L., & Fraysse, B. (2009). How much residual hearing is 'useful' for music perception with cochlear implants?. *Audiology and Neurotology*, 14(Suppl. 1), 14-21.

- Ewert, S. D. (2013). AFC—A modular framework for running psychoacoustic experiments and computational perception models. In Proceedings of the international conference on acoustics AIA-DAGA (Vol. 2013, pp. 1326-1329).
- Francart, T., & McDermott, H. J. (2013). Psychophysics, fitting, and signal processing for combined hearing aid and cochlear implant stimulation. *Ear and hearing*, 34(6), 685-700.
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* 110(2), 1150–1163. <https://doi.org/10.1121/1.1381538>
- Gaudrain E (2016) Vocoder, v1.0. Online code at <https://github.com/egaudrain/vocoder>.
<https://doi.org/10.5281/zenodo.48120>
- Gaudrain, E., Başkent, D. (2015) Factors limiting vocal-tract length discrimination in cochlear implant simulations. *J Acoust Soc Am* 137:1298–1308. <https://doi.org/10.1121/1.4908235>
- Gfeller, K., Woodworth, G., Robin, D. A., Witt, S., & Knutson, J. F. (1997). Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear and hearing*, 18(3), 252-260.
- Gfeller, K., Witt, S., Mehr, M. A., Woodworth, G., & Knutson, J. (2002). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Annals of Otology, Rhinology & Laryngology*, 111(4), 349-356.
- Gfeller, K., Oleson, J., Knutson, J. F., Breheny, P., Driscoll, V., & Olszewski, C. (2008). Multivariate predictors of music perception and appraisal by adult cochlear implant users. *Journal of the American Academy of Audiology*, 19(02), 120-134.
- Gilbert, M. L., Deroche, M. L., Jiradejvong, P., Barrett, K. C., & Limb, C. J. (2022). Cochlear implant compression optimization for musical sound quality in MED-EL users. *Ear and Hearing*, 43(3), 862-873.
- Greenwood, D.D. (1990). A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am* 87(6):2592–2605
- Grimm, G., Herzke, T., Berg, D., & Hohmann, V. (2006). The master hearing aid: A PC-based platform for algorithm development and evaluation. *Acta acustica united with Acustica*, 92(4), 618-628.
- Illg, A., Bojanowicz, M., Lesinski-Schiedat, A., Lenarz, T., & Büchner, A. (2014). Evaluation of the bimodal benefit in a large cohort of cochlear implant subjects using a contralateral hearing aid. *Otology & Neurotology*, 35(9), e240-e244.
- ITU-R – Radiocommunication Sector of ITU. 2014a. Recommendation ITU-R BS.1534-2 – Method for the subjective assessment of intermediate quality level of audio systems. BS Series, Broadcasting service (sound). International Telecommunication Union
- Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology research*, 1(1), 88-90.
- Kirchberger, M., & Russo, F. A. (2016). Dynamic range across music genres and the perception of dynamic compression in hearing-impaired listeners. *Trends in Hearing*, 20, 2331216516630549.

Landsberger, D. M., Vermeire, K., Stupak, N., Lavender, A., Neukam, J., Van de Heyning, P., & Svirsky, M. A. (2020). Music is more enjoyable with two ears, even if one of them receives a degraded signal provided by a cochlear implant. *Ear and Hearing*, 41(3), 476-490.

Landwehr, M., Fürstenberg, D., Walger, M., von Wedel, H., & Meister, H. (2014). Effects of various electrode configurations on music perception, intonation and speaker gender identification. *Cochlear Implants International*, 15(1), 27-35.

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. In *International conference on digital audio effects* (Vol. 237, p. 244).

Limb, C. J., Mo, J., Jiradejvong, P., & Jiam, N. T. (2022). The impact of vocal boost manipulations on musical sound quality for cochlear implant users. *The Laryngoscope*.

Looi, V., McDermott, H., McKay, C., & Hickson, L. (2004). Pitch discrimination and melody recognition by cochlear implant users. In *International Congress Series* (Vol. 1273, pp. 197-200). Elsevier.

Lorho, G., Le Ray, G. & Zacharov, N. 2010. eGauge – A measure of assessor expertise in audio quality evaluations. In *Audio Engineering Society Conference: 38th International Conference: Sound Quality Evaluation* (pp. 191–200). Audio Engineering Society

McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends in amplification*, 8(2), 49-82.

Moore, B.C.J. (2022) Listening to Music Through Hearing Aids: Potential Lessons for Cochlear Implants. *Trends in Hearing*; 26.

Riley, P.E., Ruhl, D.S., Camacho, M., Tolisano, A.M. (2018). Music appreciation after cochlear implantation in adult patients: a systematic review. *Otolaryngol Head Neck Surg.*, 158(6):1002-1010

Roy, A. T., Jiradejvong, P., Carver, C., & Limb, C. J. (2012). Musical sound quality impairments in cochlear implant (CI) users as a function of limited high-frequency perception. *Trends in amplification*, 16(4), 191-200.

Roy, A. T., Carver, C., Jiradejvong, P., & Limb, C. J. (2015). Musical sound quality in cochlear implant users: a comparison in bass frequency perception between fine structure processing and high-definition continuous interleaved sampling strategies. *Ear and hearing*, 36(5), 582-590.

Rubinstein, J. T. (2004). How cochlear implants encode speech. *Current opinion in otolaryngology & head and neck surgery*, 12(5), 444-448.

Scherf, F. W., Arnold, L. P., & Poster presentation at the 12th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, ESPO 2012, Amsterdam, the Netherlands, SFORL 2012, Paris, France. (2014). Exploring the clinical approach to the bimodal fitting of hearing aids and cochlear implants: Results of an international survey. *Acta oto-laryngologica*, 134(11), 1151-1157.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Lurnagaray, D., ... & Pumford, J. (2005). The desired sensation level multistage input/output algorithm. *Trends in amplification*, 9(4), 159-197.

Sucher, C. M., & McDermott, H. J. (2009). Bimodal stimulation: benefits for music perception and sound quality. *Cochlear Implants International*, 10(S1), 96-99.

Vaisberg, J. M., Beaulac, S., Glista, D., Macpherson, E. A., & Scollie, S. D. (2021). Perceived sound quality dimensions influencing frequency-gain shaping preferences for hearing aid-amplified speech and music. *Trends in Hearing*, 25, 2331216521989900.

Völker, C., Bisitz, T., Huber, R., Kollmeier, B., & Ernst, S. M. (2018). Modifications of the MUlti stimulus test with Hidden Reference and Anchor (MUSHRA) for use in audiology. *International journal of audiology*, 57(sup3), S92-S104.

Vroegop, J. L., Dingemanse, J. G., van der Schroeff, M. P., & Goedegebure, A. (2019). Comparing the effect of different hearing aid fitting methods in bimodal cochlear implant users. *American Journal of Audiology*, 28(1), 1-10.

Wong, L. L., Hickson, L., & McPherson, B. (2003). Hearing aid satisfaction: what does research from the past 20 years say? *Trends in amplification*, 7(4), 117-161.